Représentation des connaissances médicales

Dr Nicolas Jay
nicolas.jay@medecine.uhp-nancy.fr

Faculté de Médecine de Nancy

Informatique médicale – L2 Médecine 2011–2012



Plan



Qu'est-ce que l'informatique médicale ?

Définition (Greenes and Shortliffe)

Le domaine concerné par les tâches de raisonnement, de traitement de l'information, de communication dans la pratique, l'enseignement et la recherche médicales, y compris les sciences de l'information et les technologies qui supportent ces tâches.



Qu’est-ce que l’informatique médicale ?

Définition (Shortlife and Blois)

le domaine scientifique qui s”occupe de l’information biomédicale, des données et de la connaissance : leur collecte, leur extraction et leur utilisation optimale pour résoudre des problèmes et faciliter la décision.



Qu’est-ce que l’informatique médicale ?

Définition (Van Bemmel)

… comprends les aspects théoriques et pratiques du traitement de l’information et de la communication, basés sur la connaissance et l’expérience dérivées des processus en médecine et en santé.



Qu'est-ce que l'informatique médicale ?

Définition (Musen and van Bemmel)

En informatique médicale, nous développons et validons des méthodes et des systèmes pour l'acquisition, le traitement, et l'interprétation des données des patients avec l'aide de la connaissance issue de la recherche scientifique.



Alors ?

- données, information, connaissance
- traitement, communication, interprétation, raisonnement
- méthodes, systèmes, technologies



Données, information, connaissance

Une mère prend la température de son bébé, Kévin, avec un thermomètre tympanique : elle lit $38{,}5°\mathrm{C}$

- $38{,}5°\mathrm{C}$: donnée
- température tympanique = $38{,}5°\mathrm{C}$: information
- température tympanique $> 38°\mathrm{C}$: information
- température tympanique $> 38°\mathrm{C}$ donc fièvre : connaissance



Et si la mère était une machine ?

- acquisition de la donnée $D1$: $38{,}5$
- type de donnée : décimal
- unité : $^\circ\mathrm{C}$
- définition de la donnée : c'est une température
- acquisition de la donnée $D2$: tympanique
- acquisition de la donnée $D3$: mesurée chez le bébé, Kévin
- connaissance C : température tympanique > $> 38^\circ\mathrm{C}$ donc fièvre
- méthode M : comparer la température à un seuil $> 38{,}5 > 38?$
- raisonnement : utiliser $D1{,}D2{,}D3{,}M$ et C pour conclure
- communication : «Kévin a de la fièvre»



Plan



De la nécessité de standardiser les connaissances médicales

- Le système de soins produits des quantités phénoménales de données
- cliniques : compte-rendu d’hospitalisation, compte-rendu opératoire, résultats d’examens (biologie, imagerie, paracliniques), dossier infimier
- décisionnelles : recommandations, guide de bonnes pratiques, littérature scientifique
- médico-éconmiques : SIH, PMSI
- de veille sanitaire, d’épidémiologie : registres, sentinelles
- la plupart du temps informatisées



De la nécessité de standardiser les connaissances médicales

- comment réutiliser au mieux ces données pour
- améliorer la prise en charge des patients au niveau individuel,
- améliorer le fonctionnement global du système de soins
- faciliter la recherche
- automatiser certaines taches : télémédecine, monitoring
- faciliter les échanges d'information



De la nécessité de standardiser les connaissances médicales

- Le plus souvent, les «faits» médicaux sont sous forme non structurée : textes en langage naturel, images...
- pour pouvoir traiter, échanger de l'information et produire de la connaissance, une machine utilise un formalisme :
  - un lexique : ensemble de termes
  - une syntaxe : manière d'assembler les termes pour former des phrases
  - un sens : l'interprétation à donner à une phrase



Standardisation des concepts médicaux

- une machine doit pouvoir au minimum comparer des concepts
  - identiques
  - partiellement identiques
  - différents
  - exemple : myocarde et cœur
- inventaire des concepts d'un domaine (ex : la cardiologie)
- inventaire des relations entre ces concepts :
  - le myocarde est une partie du cœur
  - le myocarde est une sorte de muscle



Difficultés de standardisation

- manque de consensus : fièvre
  - OMS : T > 37,5
  - Brighton group : T > 38
- polysémie :
  - genou de la coronaire droite
  - ostéotomie du genou droit
- imprécision : infarctus
  - sylvien
  - du myocarde
  - mésentérique
- synonymie
  - pelvispondylite rhumatismale
  - spondylarthrite ankylosante



Difficultés de standardisation

- composition : si la machine connaît
  - fracture
  - fémorale, fémur
- fracture fémorale = fracture de l'os du fémur
- opération complexe pour une machine
- artère fémorale ?
- solutions :
  - approches terminologiques
  - approches compositionnelles

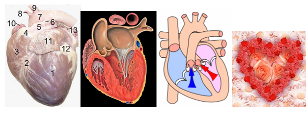


Plan



Approches terminologiques

Vision du monde fondée sur

- des objets : 

- des concepts : l'idée de cœur
- des termes : le mot «cœur»



Terminologie

Définition

recensement des concepts d’un domaine et des termes qui les désignent

- un concept
- un code unique
- un terme de référence
- une définition
- des relations généralement simples (généralisation, spécialisation)
- des synonymes
- des traductions (terminologies multilingues)
- ne gèrent pas la composition



Exemple : le Medical Subject Headings

MeSH Heading	Orbit
Tree Number	A02.835.232.781.324.690
Scope Note	Bony cavity that holds the eyeball and its associated tissues and appendages.
Entry Term	Eye Socket
Date of Entry	19990101
Unique ID	D009915

MeSH
- [A] Anatomie
  - Appareil locomoteur
    - Squelette
      - Os et tissu osseux
        - Crâne
          - Os de la face
            - Os zygomatique
            - Vomer
            - Cornets
            - **Orbite**
            - Os nasal
            - Mâchoire



Différents objectifs, différents types de terminologies

Recherche d'information : thésaurus

- phase d'indexation de documents : annotation des documents avec les termes d'un thésaurus
- phase de recherche de documents
  - mise en relation des termes d'une requête avec les termes d'indexation
  - extraction des documents correspondants

Exemple

- base documentaire : pubmed
- thésaurus : MeSH



Différents objectifs, différents types de terminologies

Recueil orienté de données : classification

- on veut classer des situations dans un but précis d'analyse (enquête épidémiologique, médico-économique)
- dans la terminologie, les classes sont construites dans ce but : partitionement équilibré des réponses possibles
- granularité variable : idéalement, plusieurs niveaux dans une même classification hiérarchique
- un concept = une classe

Exemple

- Classification statistique internationale des maladies et des problèmes de santé connexes (CIM)
- conçue pour une analyse statistique des maladies



Différents objectifs, différents types de terminologies

Recueil ouvert de données : nomenclature

- le but est de décrire des informations cliniques le plus précisément et fidèlement possible, sans objectif précis
- une nomenclature recense tous les concepts d'un domaine
- granularité la plus fine possible
- pas de regroupement d'entités pour une même entrée

Exemple

- Classification Commune des Actes Médicaux (CCAM)
- Catalogue des Actes de Rééducation Réadaptation (CdARR)



La SNOMED Internationale , une nomenclature multi-axiale

- Systematized Nomenclature of Medicine
- huit axes principaux :
  - Topographie (T)
  - Morphologie (M)
  - Fonction (F)
  - Organismes vivants (L)
  - Médicaments, produits chimiques et biologiques (C)
  - Agents, activités physiques et forces naturelles (A)
  - Métiers et professions (J)
  - Contexte social (S)
- chaque axe est une hiérarchie : « brûlure » (M-11100) est une sorte de « blessure thermique » (M-11000) qui est une sorte de « blessure » (M-10000)



La SNOMED Internationale , une nomenclature multi-axiale

- + un axe Qualificatifs et termes relationnels (G)
- + deux classifications, une pour les actes (P) et une pour les diagnostics (D : CIM-9)
- 11 axes au total
- un pas vers la composition : l'axe (G) contient des concepts supplémentaires servant à qualifier ces concepts ou à préciser leurs liens dans le concept complexe.



La SNOMED Internationale , un pas vers la composition

Exemple

Cervicalgie chez un patient affecté par une gastro-entérite. Gastro entérite avec douleur abdominale. Brève perte de connaissance sans mouvement anormal, sans douleur thoracique, sans palpitations.

F-16250	cervicalgie
G-0126	patient affecté
D5-41706	gastro-entérite
G-C008	avec
F-50860	douleur abdominale
G-C008	avec
F-A5574	brève perte de connaissance
G-C009	sans
A-80450	mouvement
G-A210	anormal
G C009	sans
F-37000	douleur thoracique
G-C009	sans
F-37150	palpitations

Plan



Approches compositionnelles

Limites des terminologies : difficultés à représenter des concepts complexes à partir de concepts simples

- nécessité de créer un terme pour chaque objet : approche pré-coordonnée
- recours à des opérateurs booléens OU, ET, NON. Ex : pneumopathie ET LobeInférieurDroit ET (NON Bacterie)
- axes orthogonaux : approche post-coordonnée
  - appendicite aiguë en SNOMED Internationale
  - appendice vermiculaire : T-59200
  - aigu : G-A231
  - inflammation : M-41000
- 6 000 termes de morphologie x 13 000 termes de topographies = 78 millions de concepts
- CIM10 : 10 000 diagnostics



Création de relations explicites

Patient de 50 ans, diabétique, hospitalisé au CHU de Nancy pour douleur thoracique.

- Homme ET
- Age=50 ET
- diabète ET
- hospitalisé ET
- CHU de Nancy ET
- Douleur thoracique



Création de relations explicites

Patient de 50 ans, diabétique, hospitalisé au CHU de Nancy pour douleur thoracique aiguë.

- Patient de sexe masculin
- Patient âgé de 50 ans
- Patient atteint de diabète
- Patient hospitalisé au CHU de Nancy
- Patient hospitalisé pour douleur thoracique
- Douleur thoracique de type aiguë

Même si la SNOMED Internationale contient des relations (axe G), elle ne donne pas de mode d'emploi, ni de contraintes



Contraintes sur les relations

- appendicite = inflammation localisée dans l'appendice
- Neurologie = service localisé dans Hôpital central
- appendicite localisé dans Hôpital central = ?
- {morphologie} affectant {localisation anatomique}
- {service} situé dans {établissement géographique}



Systèmes de représentation de connaissance

Définition

- une hiérarchie de concepts élémentaires
- une hiérarchie de relations (rôles)
- des contraintes (restrictions) sur les rôles
- permettant de fabriquer des concepts complexes par composition

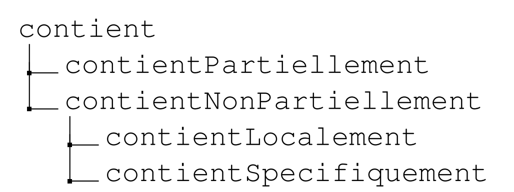


Exemple d'une hiérarchie de concepts dans OpenGALEN

```
ProcessusAnormal
 └─ ProcessusPathologique
     ├─ ProcessusDeCroissanceAnormale
     ├─ ProcessusEmbolique
     ├─ ...
     └─ AnomalieNommée
         └─ Acne
             ├─ AcneFulminans
             ├─ AcneRosace
             └─ ...
```



Exemple d'une hiérarchie de rôles dans OpenGALEN



Exemple

`CortexCérébral contientPartiellement deLaMatièreGrise`
`CavitéAbdominale contientSpecifiquement AorteAbdominale`
==> `CavitéAbdominale contientNonPartiellement AorteAbdominale`



Exemple de fabrication d'un concept complexe dans OpenGALEN

OedemaOfTheFace =

- BodyStructure
  - hasSpecificLocation SkinOfFace
  - hasUniqueAssociatedProcess OedemaProcess
  - hasPathologicalStatus pathological

rôles , concepts



Subsomption

Définition (Subsomption)

un concept A est subsumé par un concept B si A est plus spécifique que B

Exemple

- $\texttt{AcneRosace} \sqsubseteq \texttt{Acne}$
- $\texttt{OedemaOfTheFace} \sqsubseteq \texttt{hasSpecificLocation}$ $\texttt{SkinOfFace}$



Classification

- La subsomption est un élément clé pour raisonner avec un système à base de connaissance
- elle permet de comparer des concepts complexes
- pour faire de la classification : recherche des classes auxquelles appartient un objet

Exemple

- L'oedème facial est-il un processus pathologique ? $\Leftrightarrow$ ceci est-il VRAI ?
  - $\texttt{OedemaOfTheFace} \sqsubseteq \texttt{hasPathologicalStatus pathological}$
- Quels sont les maladies qui ont pour symptôme la malabsorption ? $\Leftrightarrow$ on recherche les concepts $\texttt{X}$ tels que :
  - $\texttt{X} \sqsubseteq \texttt{hasSymptom Malabsorption}$



Classification

Exemple

On veut rechercher dans des dossier électroniques tous les patients qui présentent un oedème cutané :

```
CutaneousOedema =
 BodyStructure
  |__hasSpecificLocation Skin
  |__hasUniqueAssociatedProcess OedemaProcess
  |__hasPathologicalStatus pathological
```



Subsomption et classification

L'oedème du visage est-il un oedème cutané ?

```
CutaneousOedema =
   BodyStructure
    |__hasSpecificLocation Skin
    |__hasUniqueAssociatedProcess OedemaProcess
    |__hasPathologicalStatus pathological
OedemaOfTheFace =
   BodyStructure
    |__hasSpecificLocation SkinOfFace
    |__hasUniqueAssociatedProcess OedemaProcess
    |__hasPathologicalStatus pathological
```



Subsomption et classification

L'oedème du visage est-il un oedème cutané ?

```
CutaneousOedema =
   BodyStructure
    |__hasSpecificLocation Skin
    |__hasUniqueAssociatedProcess OedemaProcess
    |__hasPathologicalStatus pathological
OedemaOfTheFace =
   BodyStructure
    |__hasSpecificLocation SkinOfFace
    |__hasUniqueAssociatedProcess OedemaProcess
    |__hasPathologicalStatus pathological
```

- SkinOfFace $\sqsubseteq$ Skin : vrai



Subsomption et classification

L'oedème du visage est-il un oedème cutané ?

```
CutaneousOedema =
   BodyStructure
    ├─ hasSpecificLocation Skin
    ├─ hasUniqueAssociatedProcess OedemaProcess
    └─ hasPathologicalStatus pathological
OedemaOfTheFace =
   BodyStructure
    ├─ hasSpecificLocation SkinOfFace
    ├─ hasUniqueAssociatedProcess OedemaProcess
    └─ hasPathologicalStatus pathological
```

- SkinOfFace $\sqsubseteq$ Skin : vrai
- hasSpecificLocation SkinOfFace $\sqsubseteq$ hasSpecificLocation Skin : vrai



Subsomption et classification

L'oedème du visage est-il un oedème cutané ?

```
CutaneousOedema =
   BodyStructure
    |__hasSpecificLocation Skin
    |__hasUniqueAssociatedProcess OedemaProcess
    |__hasPathologicalStatus pathological
OedemaOfTheFace =
   BodyStructure
    |__hasSpecificLocation SkinOfFace
    |__hasUniqueAssociatedProcess OedemaProcess
    |__hasPathologicalStatus pathological
```

- SkinOfFace $\sqsubseteq$ Skin : vrai
- hasSpecificLocation SkinOfFace $\sqsubseteq$ hasSpecificLocation Skin : vrai
- OedemaOfTheFace $\sqsubseteq$ CutaneousOedema : vrai



Subsomption et classification

L'oedème du poumon est-il un oedème cutané ?

```
CutaneousOedema =
  BodyStructure
   |__hasSpecificLocation Skin
   |__hasUniqueAssociatedProcess OedemaProcess
   |__hasPathologicalStatus pathological
 PulmonaryOedema =
  BodyStructure
   |__hasSpecificLocation Lung
   |__hasUniqueAssociatedProcess OedemaProcess
   |__hasPathologicalStatus pathological
```



Subsomption et classification

L'oedème du poumon est-il un oedème cutané ?

```
CutaneousOedema =
  BodyStructure
   |__hasSpecificLocation Skin
   |__hasUniqueAssociatedProcess OedemaProcess
   |__hasPathologicalStatus pathological
 PulmonaryOedema =
  BodyStructure
   |__hasSpecificLocation Lung
   |__hasUniqueAssociatedProcess OedemaProcess
   |__hasPathologicalStatus pathological
```

- Lung $\sqsubseteq$ Skin : faux



Subsomption et classification

L'oedème du poumon est-il un oedème cutané ?

```
CutaneousOedema =
  BodyStructure
  |__hasSpecificLocation Skin
  |__hasUniqueAssociatedProcess OedemaProcess
  |__hasPathologicalStatus pathological
 PulmonaryOedema =
  BodyStructure
  |__hasSpecificLocation Lung
  |__hasUniqueAssociatedProcess OedemaProcess
  |__hasPathologicalStatus pathological
```

- Lung $\sqsubseteq$ Skin : faux
- hasSpecificLocation Lung $\sqsubseteq$ hasSpecificLocation Skin : faux



Subsomption et classification

L'oedème du poumon est-il un oedème cutané ?

```
CutaneousOedema =
  BodyStructure
   |__hasSpecificLocation Skin
   |__hasUniqueAssociatedProcess OedemaProcess
   |__hasPathologicalStatus pathological
 PulmonaryOedema =
  BodyStructure
   |__hasSpecificLocation Lung
   |__hasUniqueAssociatedProcess OedemaProcess
   |__hasPathologicalStatus pathological
```

- Lung $\sqsubseteq$ Skin : faux
- hasSpecificLocation Lung $\sqsubseteq$ hasSpecificLocation Skin : faux
- PulmonaryOedema $\sqsubseteq$ CutaneousOedema : faux



Subsomption et classification

Si on voulait faire la même chose avec un moteur de recherche basé sur le texte ?

- «oedème cutané» ne permet pas de retrouver «oedème de la face»
- «oedème» ramène «oedème du poumon», «oedème laryngé» ...
- « Le patient ne présente pas d'oedème cutané »

Si on voulait faire la même chose avec un moteur de recherche basé sur une terminologie ?

- moins évident
- il faut que tous les concepts se rapportant aux oedèmes cutanés existent dans la terminologie
- il faut connaître tous les termes se rapportant à ces concepts



Ontologies

Définition

Une ontologie est une spécification explicite d'une conceptualisation

- un ensemble de concepts,
- un ensemble de relations et contraintes
- effectivement utilisés pour modéliserun domaine donné.



Ontologies

- les classifications, nomenclatures, thésauri ne sont pas des ontologies
- les relations hiérarchiques entre concepts ne sont pas spécifiée
- exemple : pas de distinction entre «est-un» et «est-une-partie-de»

Exemple (MeSH)

- Head
  - Ear
  - Face
  - Scalp
  - Skull
  - Base
- Blood Cells
  - Blood Platelets
  - Erythrocytes
  - Hemocytes



Exemple de formalisme pour représenter une ontologie : les logiques de description

Une logique de description est un formalisme, en deux parties

- les informations terminologiques TBox : définition des notions basiques ou dérivées et de comment elles sont reliées entre elles. Ces informations sont "génériques" ou "globales", vraies dans tous les modèles et pour tous les individus.
- les informations sur les individus ABox : ces informations sont "spécifiques" ou "locales", vraies pour certains individus particuliers.



Exemple de formalisme pour représenter une ontologie : les logiques de description

Exemple (TBox)

- `Homme := Personne ET Male`
- `Femme := Personne ET Femelle`
- `Père := Homme aUnEnfant.Personne`
- `Mère := Femme aUnEnfant.Personne`
- `Parent := Père OU Mère`
- `GrandPère := Homme aUnEnfant.Parent`
- `GrandMère :=?`



Exemple de formalisme pour représenter une ontologie : les logiques de description

Exemple (TBox)

- `Homme := Personne ET Male`
- `Femme := Personne ET Femelle`
- `Père := Homme aUnEnfant.Personne`
- `Mère := Femme aUnEnfant.Personne`
- `Parent := Père OU Mère`
- `GrandPère := Homme aUnEnfant.Parent`
- `GrandMère :=?`
- `GrandMère := Femme aUnEnfant.Parent`



Exemple de formalisme pour représenter une ontologie : les logiques de description

Exemple (ABox)

- `Père(Pierre)` : `Pierre` est une instance de `Homme`
- `aUnEnfant(Jeanne,Pierre)` : `Jeanne` a un enfant qui est `Pierre`
- `Femme(Jeanne)` : `Jeanne` est une instance de `Femme`



Exemple de formalisme pour représenter une ontologie : les logiques de description

Exemple (Raisonnement avec une TBox et une ABox)

- Homme := Personne ET Male
- Femme := Personne ET Femelle
- Père := Homme aUnEnfant.Personne
- Mère := Femme aUnEnfant.Personne
- Parent := Père OU Mère
- GrandPère := Homme aUnEnfant.Parent
- GrandMère := Femme aUnEnfant.Parent

- Père(Pierre)
- aUnEnfant(Jeanne,Pierre)
- Femme(Jeanne)
- on en déduit : ?



Exemple de formalisme pour représenter une ontologie : les logiques de description

Exemple (Raisonnement avec une TBox et une ABox)

- Homme := Personne ET Male
- Femme := Personne ET Femelle
- Père := Homme aUnEnfant.Personne
- Mère := Femme aUnEnfant.Personne
- Parent := Père OU Mère
- GrandPère := Homme aUnEnfant.Parent
- GrandMère := Femme aUnEnfant.Parent

- Père(Pierre)
- aUnEnfant(Jeanne,Pierre)
- Femme(Jeanne)
- on en déduit : ?
  - Parent(Pierre)



Exemple de formalisme pour représenter une ontologie : les logiques de description

Exemple (Raisonnement avec une TBox et une ABox)

- Homme := Personne ET Male
- Femme := Personne ET Femelle
- Père := Homme aUnEnfant.Personne
- Mère := Femme aUnEnfant.Personne
- Parent := Père OU Mère
- GrandPère := Homme aUnEnfant.Parent
- GrandMère := Femme aUnEnfant.Parent

- Père(Pierre)
- aUnEnfant(Jeanne,Pierre)
- Femme(Jeanne)
- on en déduit : ?
  - Parent(Pierre)
  - GrandMère(Jeanne)



Plan



Utilisation des terminologies et des ontologies en médecine

- Gestion des connaissances
- Intégration, échange de données et interopérabilité sémantique
- Aide à la décision



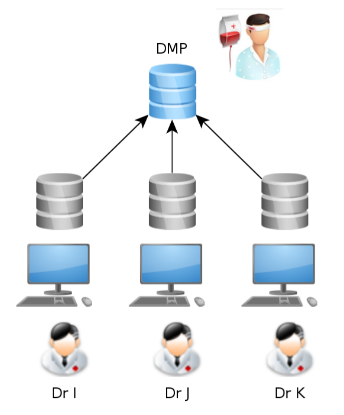
Gestion des connaissances

- Annotation, indexation :
  - manuelle ou automatique
  - articles scientifiques (PubMed), documents médicaux, gènes, protéines
- Recherche d'information :
  - requêtes avec vocabulaire contrôlé (ex : MeSH et Medline)
  - gestion des synonymes
  - traduction
  - classification de documents
- Mapping(mise en correspondance) entre différentes terminologies



Intégration, échange de données et interopérabilité sémantique

- échange de données entre logiciels différents.
  Ex : LOINC (Logical Observation Identifiers Names and Codes) transmission des résultats de laboratoire
- intégration de données issues de systèmes différents





Aide à la décision, découverte de connaissance

- aide à la décision : système d'alerte pour les allergies médicamenteuses
- traitement du langage naturel :
  - extraction d'information dans des compte rendus médicaux
  - fouille de texte : veille scientifique, découverte et synthèse de connaissances
- fouille de données : découverte automatique de liens intéressants dans les bases de données

